

# Large and moderate deviations for the left random walk on $GL_d(\mathbb{R})$

Christophe Cuny<sup>a</sup>, Jérôme Dedecker<sup>b</sup> and Florence Merlevède<sup>c</sup>

## Abstract

Using martingale methods, we obtain some upper bounds for large and moderate deviations of products of independent and identically distributed elements of  $GL_d(\mathbb{R})$ . We investigate all the possible moment conditions, from super-exponential moments to weak moments of order  $p > 1$ , to get a complete picture of the situation. We also prove a moderate deviation principle under an appropriate tail condition.

<sup>a</sup> Université de la Nouvelle-Calédonie, Equipe ERIM, et CentraleSupélec, Laboratoire MAS.  
Email: cuny@univ-nc.nc

<sup>b</sup> Université Paris Descartes, Sorbonne Paris Cité, Laboratoire MAP5 (UMR 8145).  
Email: jerome.dedecker@parisdescartes.fr

<sup>c</sup> Université Paris-Est, LAMA (UMR 8050), UPEM, CNRS, UPEC.  
Email: florence.merlevede@u-pem.fr

## 1 Introduction

Let  $(\Omega, \mathcal{F}, \mathbb{P})$  be a probability space and  $(Y_n)_{n \geq 1}$  be independent and identically distributed random variables on  $(\Omega, \mathcal{F}, \mathbb{P})$  taking values in  $G := GL_d(\mathbb{R})$ ,  $d \geq 2$  (the group of invertible  $d$ -dimensional real matrices), with common distribution  $\mu$ . Denote by  $\Gamma_\mu$  the closed semi-group generated by the support of  $\mu$ . Let  $\|\cdot\|$  be the euclidean norm on  $\mathbb{R}^d$ , and for every  $g \in GL_d(\mathbb{R})$ , let  $\|g\| := \sup_{\|x\|=1} \|gx\|$ .

In all the paper, we assume that  $\mu$  is *strongly irreducible*, i.e. that no proper finite union of subspaces of  $\mathbb{R}^d$  are invariant by  $\Gamma_\mu$  and that it is *proximal*, i.e. that there exists a matrix in  $\Gamma_\mu$  admitting a unique (with multiplicity one) eigenvalue with maximum modulus.

For such a measure  $\mu$ , it is known that there exists a unique invariant measure  $\nu$  on the projective space  $X := P_{d-1}(\mathbb{R})$  (see for instance Theorem 3.1 of [5]) in the following sense: for any bounded Borel function  $h$  from  $X$  to  $\mathbb{R}$

$$\int_X h(u) \nu(du) = \int_G \int_X h(g \cdot u) \mu(dg) \nu(du). \quad (1.1)$$

Moreover, if

$$\int_G \log N(g) \mu(dg) < \infty, \quad \text{where } N(g) := \max(\|g\|, \|g^{-1}\|), \quad (1.2)$$

then (see for instance Corollary 3.4 page 54 of [5] or Theorem 3.28 of [3]), for every  $x \in S^{d-1}$ ,

$$\frac{\log \|Y_n \cdots Y_1 x\|}{n} \xrightarrow{n \rightarrow +\infty} \lambda_\mu = \int_G \int_X \sigma(g, u) \mu(dg) \nu(du) \quad \text{almost surely,} \quad (1.3)$$

where

$$\sigma(g, \bar{x}) = \log \left( \frac{\|g \cdot x\|}{\|x\|} \right) \quad \text{for } g \in GL_d(\mathbb{R}) \text{ and } x \in \mathbb{R}^d - \{0\}, \bar{x} \text{ denoting the class of } x \text{ in } X.$$

Note that the function  $\sigma$  defined above is a *cocycle*, in the following sense:

$$\sigma(gg', u) = \sigma(g, g' \cdot u) + \sigma(g', u) \quad \text{for any } g, g' \in G \text{ and } u \in X. \quad (1.4)$$

Let  $A_k = Y_k \cdots Y_1$  for  $k \geq 1$  and  $A_0 = \text{Id}$ . In this paper we wish to study the asymptotic behavior of

$$\sup_{\|x\|=1} \mathbb{P} \left( \max_{1 \leq k \leq n} |\log \|A_k x\| - k \lambda_\mu| > n^\alpha y \right), \quad (1.5)$$

when  $\alpha \in (1/2, 1]$ , under stronger moment conditions on  $\log N(Y_1)$  than (1.2). This is a way to study rates of convergence in the strong law (1.3). In the probabilistic terminology, the case  $\alpha \in (1/2, 1)$  corresponds to the *moderate deviation* regime, and  $\alpha = 1$  to the *large deviation* regime.

The case  $\alpha = 1/2$  corresponds to the normalization of the central limit theorem. In that case, the asymptotic behavior of (1.5) is due to Benoist and Quint [4] as soon as  $\log N(Y_1)$  has a moment of order 2 (note that Benoist and Quint do not deal with the maximum in (1.5), but their method also applies in that case, see also Theorem 1(ii) in [6]). A previous result is due to Jan [18] under a moment of order  $2 + \epsilon$ ,  $\epsilon > 0$ .

In this paper, we shall give precise informations on the rate of convergence to 0 (as  $n \rightarrow \infty$ ) of (1.5) when  $\alpha \in (1/2, 1]$ , under various moment conditions on the random variable  $\log N(Y_1)$ : sub or super-exponential moments in Section 2, weak moments of order  $p > 1$  in Section 3, and strong moments of order  $p \geq 1$  in Section 4.

In Section 2.2 we shall give a moderate deviation principle for the process

$$\{n^{-\alpha}(\log \|A_{[nt]} x\| - [nt] \lambda_\mu), t \in [0, 1]\}$$

when  $\log N(Y_1)$  satisfies Arcones's tail condition [1] (which is true under an appropriate sub-exponential moment condition). In Section 4 we obtain some results in the spirit of Baum and Katz [2] which complement the results on complete convergence obtained in [4] in the case  $\alpha = 1$ . When  $\log N(Y_1)$  has a strong moment of order  $p \in (1, 2)$  and  $\alpha = 1/p$ , this gives the rate  $n^{(p-1)/p}$  in the strong law of large numbers, which was proved by another method in [6], Theorem 1(i).

All along the paper, the following notations will be used: let  $\mathcal{F}_0 = \{\emptyset, \Omega\}$  and  $\mathcal{F}_k = \sigma(Y_1, \dots, Y_k)$ , for any  $k \geq 1$ . For any  $x \in S^{d-1}$ , define  $X_{0,x} = x$  and  $X_{n,x} = \sigma(Y_n, A_{n-1}x)$  for  $n \geq 1$ . With these notations, for any  $x \in S^{d-1}$  and any positive integer  $k$ ,

$$\log \|A_k x\| = \sum_{i=1}^k X_{i,x}. \quad (1.6)$$

The equality (1.6) follows easily from the fact that  $\sigma$  is a cocycle (*i.e.* (1.4) holds). In Section 6 we shall present some extensions of our results to general cocycles, in the spirit of Benoist and Quint [4].

## 2 The case of (sub/super) exponential moments

### 2.1 Upper bounds for large deviations

Let  $r > 0$ . In this subsection, we assume that

$$\int_G e^{\delta(\log N(g))^r} \mu(dg) < \infty, \quad \text{for some } \delta > 0. \quad (2.1)$$

We first consider the case  $r \geq 1$ . In that case, using the spectral gap property, Le Page [19] proved the following large deviation principle: there exists a positive constant  $A$  such that, for any  $y \in (0, A)$ ,

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{P}(|\log \|A_n x\| - n\lambda_\mu| > ny) = \phi(y). \quad (2.2)$$

Of course, this is the best possible result for  $y \in (0, A)$ . However, it does not give any information for large values of  $y$ , and the rate function  $\phi$  is not explicit (in particular, one cannot easily describe the behavior of  $\phi$  when  $r$  varies in  $[1, \infty)$ ).

The following result, which is obtained *via* a completely different method, can be seen as a complementary result of (2.2). It gives an explicit (up to a constant) upper bound for  $\phi$  when  $y \in (0, A)$ , and this upper bound is valid for any  $y > 0$ . In particular, we can see the qualitative change in the behavior of large deviations for large  $y$  when  $r$  varies in  $[1, \infty)$ .

**Theorem 2.1.** *Assume that (2.1) holds for some  $r \geq 1$ . Then there exists a positive constant  $C$  such that, for any  $y > 0$ ,*

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \log \sup_{\|x\|=1} \mathbb{P} \left( \max_{1 \leq k \leq n} |\log \|A_k x\| - k\lambda_\mu| > ny \right) \leq -C (y^2 \mathbf{1}_{y \in (0,1)} + y^r \mathbf{1}_{y \geq 1}). \quad (2.3)$$

For  $r \in (0, 1)$ , there is no such result as (2.2). Instead, one can prove:

**Theorem 2.2.** *Assume that (2.1) holds for some  $r \in (0, 1)$ . Then there exists a positive constant  $C$  such that, for any  $y > 0$ ,*

$$\limsup_{n \rightarrow \infty} \frac{1}{n^r} \log \sup_{\|x\|=1} \mathbb{P} \left( \max_{1 \leq k \leq n} |\log \|A_k x\| - k\lambda_\mu| > ny \right) \leq -C y^r. \quad (2.4)$$

**Proofs of Theorems 2.1 and 2.2.** Since  $\int (\log N(g))^2 \mu(dg) < \infty$ , we infer from the equality (3.9) in [4] that, for any  $x \in S^{d-1}$ ,

$$X_{k,x} - \lambda_\mu = D_{k,x} + \psi(A_{k-1}x) - \psi(A_kx), \quad (2.5)$$

where  $\psi$  is a bounded function and  $D_{k,x}$  is  $\mathcal{F}_k$ -measurable and such that  $\mathbb{E}(D_{k,x}|\mathcal{F}_{k-1}) = 0$ . The decomposition (2.5) is called a *martingale-coboundary* decomposition. Such a decomposition has been used for the first time in the paper [13] by Gordin (see also [14]).

Starting from (1.6) and (2.5), for any  $x \in S^{d-1}$  and any positive integer  $k$ ,

$$\log \|A_kx\| - k\lambda_\mu = M_{k,x} + \psi(x) - \psi(A_kx), \quad (2.6)$$

where  $M_k(x) = D_{1,x} + \dots + D_{k,x}$  is a martingale adapted to the filtration  $\mathcal{F}_k$ . Clearly, since  $|\psi(x) - \psi(A_kx)| \leq 2\|\psi\|_\infty$ , it is equivalent to prove (2.3) and (2.4) for  $M_{k,x}$  instead of  $(\log \|A_kx\| - k\lambda_\mu)$ .

To do this, we first note that if (2.1) holds, then

$$\left\| \mathbb{E} \left( e^{\delta |X_{k,x}|^r} \middle| \mathcal{F}_{k-1} \right) \right\|_\infty = \left\| \int e^{\delta |\sigma(g, A_{k-1}x)|^r} \mu(dg) \right\|_\infty \leq \int_G e^{\delta (\log N(g))^r} \mu(dg) < \infty. \quad (2.7)$$

Using again that  $\psi$  is bounded we infer from (2.5) and (2.7) that there exists a constant  $K$  such that

$$\sup_{\|x\|=1} \left\| \mathbb{E} \left( e^{\delta |D_{k,x}|^r} \middle| \mathcal{F}_{k-1} \right) \right\|_\infty < K, \quad (2.8)$$

for any positive integer  $k$ .

Starting from (2.8), it remains to apply known results to the martingale  $M_k(x)$ .

To prove (2.3) (case  $r \geq 1$ ), we apply Theorem 1.1 of [20], which implies that there exists a positive constant  $c$  such that, for any  $y > 0$ ,

$$\sup_{\|x\|=1} \mathbb{P} \left( \max_{1 \leq k \leq n} |M_{k,x}| > ny \right) \leq 2 \exp \left( -nc \left( y^2 \mathbf{1}_{y \in (0,1)} + y^r \mathbf{1}_{y \geq 1} \right) \right). \quad (2.9)$$

The upper bound (2.3) follows directly from (2.9). Note that a direct application of Theorem 1.1 of [20] gives us (2.9) without the maximum over  $k$ . However, a careful reading of the proof reveals that one can take the maximum over  $k$ . The only argument that should be added to the proof is Doob's maximal inequality for non-negative submartingales, which implies that

$$\mathbb{E} \left( e^{\lambda \max_{1 \leq k \leq n} M_{k,x}} \right) \leq \mathbb{E} \left( e^{\lambda M_{n,x}} \right), \text{ for any } \lambda > 0.$$

To prove (2.4) (case  $r \in (0,1)$ ), we apply Theorem 2.1 of [12] (see also the proof of Proposition 3.5 of [7]) and more precisely the upper bound (13) in [12], which implies that there exist two positive constants  $c_1$  and  $c_2$  such that

$$\sup_{\|x\|=1} \mathbb{P} \left( \max_{1 \leq k \leq n} |M_{k,x}| > ny \right) \leq 4 \exp \left( -c_1 (ny)^r \right), \text{ for any } y > c_2 n^{-(1-r)/(2-r)}. \quad (2.10)$$

The upper bound (2.4) follows directly from (2.10).  $\diamond$

## 2.2 A moderate deviation principle

Let  $(b_n)_{n \geq 0}$  be a sequence of positive numbers satisfying the following regularity conditions:

$$\text{The functions } f(n) = \frac{n^2}{b_n^2} \text{ and } g(n) = b_n^2 \text{ are strictly increasing to infinity, and } \lim_{n \rightarrow \infty} \frac{n}{b_n^2} = 0. \quad (2.11)$$

For  $x \in S^{d-1}$ , let

$$Z_{n,x} = \left\{ \frac{\log \|A_{[nt]}x\| - [nt]\lambda_\mu}{b_n}, t \in [0, 1] \right\}.$$

The process  $Z_{n,x}$  takes values in the space  $D([0, 1])$  equipped with the usual Skorokhod topology. The following functional moderate deviation principle holds:

**Theorem 2.3.** *Let  $(b_n)_{n \geq 0}$  be a sequence of positive numbers satisfying the condition (2.11). Assume  $\int (\log N(g))^2 \mu(dg) < \infty$  and*

$$\limsup_{n \rightarrow \infty} \frac{n}{b_n^2} \log n \mu \{ \log N > b_n \} = -\infty. \quad (2.12)$$

*Then, for any  $x \in S^{d-1}$ ,  $n^{-1} \mathbb{E}((\log \|A_n x\| - n\lambda_\mu)^2) \rightarrow V$  as  $n \rightarrow \infty$ , where  $V$  does not depend on  $x$ . Moreover, for any Borel set  $\Gamma \subset D([0, 1])$ ,*

$$-\inf_{\varphi \in \Gamma^\circ} I_V(\varphi) \leq \liminf_{n \rightarrow \infty} \frac{n}{b_n^2} \log \inf_{\|x\|=1} \mathbb{P}(Z_{n,x} \in \Gamma) \leq \limsup_{n \rightarrow \infty} \frac{n}{b_n^2} \log \sup_{\|x\|=1} \mathbb{P}(Z_{n,x} \in \Gamma) \leq -\inf_{\varphi \in \bar{\Gamma}} I_V(\varphi), \quad (2.13)$$

where

$$I_V(h) = \frac{1}{2V} \int_0^1 (h'(u))^2 du$$

if simultaneously  $V > 0$ ,  $h(0) = 0$  and  $h$  is absolutely continuous, and  $I_V(h) = +\infty$  otherwise.

**Remark 2.1.** Let  $(b_n)_{n \geq 0}$  be a sequence of positive numbers satisfying (2.11). If  $(X_i)_{i \geq 1}$  is a sequence of independent and identically distributed random variables, Arcones [1] proved that the functional moderate deviation principle holds provided  $\mathbb{E}(X_1^2) < \infty$  and

$$\limsup_{n \rightarrow \infty} \frac{n}{b_n^2} \log n \mathbb{P}(|X_1| > b_n) = -\infty. \quad (2.14)$$

Moreover, he showed that condition (2.14) is also necessary for the moderate deviation principle. Note that our condition (2.12) is exactly Arcones's tail condition for the random variable  $\log N(Y_1)$ . When  $b_n = n^\alpha$  with  $\alpha \in (1/2, 1)$ , the tail condition (2.12) is true if

$$\mu \{ \log N > x \} \leq e^{-x^\beta a(x)},$$

for  $\beta = 2 - (1/\alpha)$  and a function  $a$  such that  $a(x) \rightarrow \infty$  as  $x \rightarrow \infty$  (note that  $\beta \in (0, 1)$ , so only a sub-exponential moment is needed for  $\log N(Y_1)$ ).

**Remark 2.2.** Applying the contraction principle, Theorem 2.3 implies in particular that, for any Borel set  $\Gamma \subset \mathbb{R}^+$ ,

$$\begin{aligned} -\inf \left\{ \frac{y^2}{2V}, y \in \Gamma^\circ \right\} &\leq \liminf_{n \rightarrow \infty} \frac{n}{b_n^2} \log \inf_{\|x\|=1} \mathbb{P} \left( \frac{\max_{1 \leq k \leq n} |\log \|A_k x\| - k\lambda_\mu|}{b_n} \in \Gamma \right) \\ &\leq \limsup_{n \rightarrow \infty} \frac{n}{b_n^2} \log \sup_{\|x\|=1} \mathbb{P} \left( \frac{\max_{1 \leq k \leq n} |\log \|A_k x\| - k\lambda_\mu|}{b_n} \in \Gamma \right) \leq -\inf \left\{ \frac{y^2}{2V}, y \in \bar{\Gamma} \right\}. \end{aligned}$$

Note that a partial result in this direction has been obtained in [3], Proposition 11.12. In that Proposition, the authors proved a moderate deviation principle for  $(\log \|A_n x\| - n\lambda_\mu)$  and the collection of open intervals, under an exponential moment for  $\log N(Y_1)$ . However, their result is stated in a more general framework than ours (see Section 6 of the present paper for an extension of Theorem 2.3 to general cocycles).

**Proof of Theorem 2.3.** Since  $\int (\log N(g))^2 \mu(dg) < \infty$ , the decomposition (2.6) holds, and it is equivalent to prove (2.13) for the process

$$\tilde{Z}_{n,x} = \left\{ \frac{M_{[nt],x}}{b_n}, t \in [0, 1] \right\}$$

instead of  $Z_{n,x}$ . Now, by a standard argument, to get the result uniformly with respect to  $x \in S^{d-1}$  in (2.13), it suffices to prove the functional moderate deviation principle for the process  $\tilde{Z}_{n,x_n}$ , where  $(x_n)_{n \geq 1}$  is any sequence of points in  $S^{d-1}$ .

The result will follow from the next proposition, which is a triangular version of Theorem 1 in [11]. This proposition is in fact a corollary of a more general result for triangular arrays of martingale differences which can be deduced from Puhalskii's results and Worms's paper (see [22] and [23]). We refer to Theorem 5.1 of the Appendix for a complete statement and some elements of proof.

Before giving the statement of this proposition, we need more notations. Assuming (2.11), we can construct the strictly increasing continuous function  $f(x)$  that is formed by the line segments from  $(n, f(n))$  to  $(n+1, f(n+1))$ . Similarly we define  $g(x)$  and denote by

$$c(x) = f^{-1}(g(x)). \quad (2.15)$$

**Proposition 2.4.** *Let  $(d_{i,n})_{1 \leq i \leq n}$  be a triangular array of real-valued square-integrable martingale differences, adapted to a triangular array of filtrations  $(\mathcal{F}_{i,n})_{0 \leq i \leq n}$ . Let  $(b_n)_{n \geq 0}$  be a sequence of positive numbers satisfying (2.11), and let*

$$\bar{Z}_n = \left\{ \frac{d_{1,n} + \dots + d_{[nt],n}}{b_n}, t \in [0, 1] \right\}.$$

*Assume that the three following conditions holds*

1. *There exists a positive number  $V$  such that, for any  $\delta > 0$  and any  $t \in [0, 1]$ ,*

$$\limsup_{n \rightarrow \infty} \frac{n}{b_n^2} \log \mathbb{P} \left( \left| \left( \frac{1}{n} \sum_{i=1}^{[nt]} \mathbb{E}(d_{i,n}^2 | \mathcal{F}_{i-1,n}) \right) - Vt \right| > \delta \right) = -\infty. \quad (2.16)$$

2. For any  $\varepsilon > 0$  and  $\delta > 0$ ,

$$\limsup_{n \rightarrow \infty} \frac{n}{b_n^2} \log \mathbb{P} \left( \frac{1}{n} \sum_{i=1}^n \mathbb{E} \left( d_{i,n}^2 \mathbf{1}_{|d_{i,n}| > \varepsilon n b_n^{-1}} \middle| \mathcal{F}_{i-1,n} \right) > \delta \right) = -\infty. \quad (2.17)$$

3.

$$\frac{n}{b_n^2} \log \left( \sup_{n \leq m \leq c(n+1)} \sup_{1 \leq k \leq m} n \left\| \mathbb{P}(|d_{k,m}| > b_n | \mathcal{F}_{k-1,m}) \right\|_\infty \right) \rightarrow -\infty \text{ as } n \rightarrow \infty, \quad (2.18)$$

where  $c(n)$  is defined in (2.15).

Then, for any Borel set  $\Gamma \subset D([0, 1])$ ,

$$- \inf_{\varphi \in \Gamma^\circ} I_V(\varphi) \leq \liminf_{n \rightarrow \infty} \frac{n}{b_n^2} \log \mathbb{P}(\bar{Z}_n \in \Gamma) \leq \limsup_{n \rightarrow \infty} \frac{n}{b_n^2} \log \mathbb{P}(\bar{Z}_n \in \Gamma) \leq - \inf_{\varphi \in \Gamma} I_V(\varphi), \quad (2.19)$$

where  $I_V$  is defined as in Theorem 2.3.

Let us conclude the proof of Theorem 2.3. Let  $(x_n)_{n \geq 1}$  be any sequence of points in  $S^{d-1}$ . We apply Proposition 2.4 to the martingale differences  $d_{i,n} = D_{i,x_n}$  (recall that  $D_{i,x}$  is the martingale difference of the decomposition (2.5)). Condition (2.17) is clearly satisfied thanks to (2.5) and the fact that

$$\begin{aligned} \left\| \mathbb{E} \left( X_{k,x_n}^2 \mathbf{1}_{|X_{k,x_n}| > \varepsilon n b_n^{-1}} \middle| \mathcal{F}_{k-1} \right) \right\|_\infty &= \left\| \int (\sigma(g, A_{k-1}x_n))^2 \mathbf{1}_{|\sigma(g, A_{k-1}x_n)| > \varepsilon n b_n^{-1}} \mu(dg) \right\|_\infty \\ &\leq \int_G (\log N(g))^2 \mathbf{1}_{\log N(g) > \varepsilon n b_n^{-1}} \mu(dg). \end{aligned}$$

To check Condition (2.16), we apply Proposition 3.1 in [4], which implies that, for any  $\delta > 0$  and any  $t \in [0, 1]$ , there exist  $A > 0$  and  $\alpha > 0$  such that, for the variance  $V$  defined in Theorem 2.3,

$$\mathbb{P} \left( \left| \left( \frac{1}{n} \sum_{i=1}^{[nt]} \mathbb{E}(D_{i,x_n}^2 | \mathcal{F}_{i-1}) \right) - Vt \right| > \delta \right) \leq A e^{-\alpha n}.$$

Condition (2.16) follows then easily, since  $n^2 b_n^{-2} \rightarrow \infty$  as  $n \rightarrow \infty$ .

It remains to check Condition (2.18). By (2.5) again, it is equivalent to prove the condition for  $X_{k,x_m}$  instead of  $D_{k,x_m}$ . Now

$$\left\| \mathbb{E} \left( \mathbf{1}_{|X_{k,x_m}| \geq b_n} \middle| \mathcal{F}_{k-1} \right) \right\|_\infty = \left\| \int \mathbf{1}_{\sigma(g, A_{k-1}x_m) \geq b_n} \mu(dg) \right\|_\infty \leq \mu \{ \log N \geq b_n \},$$

and the result follows by (2.12).  $\diamond$

### 3 The case of weak moment of order $p > 1$

In this section, we study the asymptotic behavior of (1.5) when  $\log N(Y_1)$  has only a weak moment of order  $p > 1$ .

**Theorem 3.1.** *Let  $p > 1$  and assume that*

$$\sup_{t>0} t^p \mu \{ \log N > t \} < \infty. \quad (3.1)$$

*Let  $\alpha \in (1/2, 1]$  and  $\alpha \geq 1/p$ . Then there exists a positive constant  $C$  such that, for any  $y > 0$ ,*

$$\limsup_{n \rightarrow \infty} n^{\alpha p - 1} \sup_{\|x\|=1} \mathbb{P} \left( \max_{1 \leq k \leq n} |\log \|A_k x\| - k \lambda_\mu| > n^\alpha y \right) \leq \frac{C}{y^p}. \quad (3.2)$$

**Proof of Theorem 3.1.**

*The case  $p > 2$ .* In that case the decomposition (2.6) holds, and it is equivalent to prove (3.2) for  $M_{k,x}$  instead of  $(\log \|A_k x\| - k \lambda_\mu)$ . To do this, we shall apply the following inequality due to Haeusler [16]: for all  $\gamma, u, v > 0$ ,

$$\begin{aligned} \mathbb{P} \left( \max_{1 \leq k \leq n} |M_{k,x}| \geq \gamma \right) &\leq \sum_{i=1}^n \mathbb{P} (|D_{i,x}| \geq u) + 2\mathbb{P} \left( \sum_{i=1}^n \mathbb{E}(D_{i,x}^2 | \mathcal{F}_{i-1}) \geq v \right) \\ &\quad + 2 \exp \left( \gamma u^{-1} (1 - \log(\gamma u v^{-1})) \right). \end{aligned} \quad (3.3)$$

Note that if (3.1) holds for  $p > 2$ , then

$$\|\mathbb{E}(X_{k,x}^2 | \mathcal{F}_{k-1})\|_\infty = \left\| \int (\sigma(g, A_{k-1}x))^2 \mu(dg) \right\|_\infty \leq \int_G (\log N(g))^2 \mu(dg) < \infty, \quad (3.4)$$

and there exists a positive constant  $C$  such that

$$\left\| \mathbb{E} \left( \mathbf{1}_{|X_{k,x}| \geq u} | \mathcal{F}_{k-1} \right) \right\|_\infty = \left\| \int \mathbf{1}_{\sigma(g, A_{k-1}x) \geq u} \mu(dg) \right\|_\infty \leq \mu \{ \log N \geq u \} \leq \frac{C}{u^p}, \quad (3.5)$$

for any  $u > 0$  and any positive integer  $k$ . Using again that  $\psi$  is bounded we infer from (2.5) and (3.4) that there exist two positive constants  $c_1, c_2$  such that

$$\sup_{\|x\|=1} \|\mathbb{E}(D_{k,x}^2 | \mathcal{F}_{k-1})\|_\infty \leq c_1, \quad (3.6)$$

$$\sup_{u>0} u^p \sup_{\|x\|=1} \mathbb{P}(|D_{k,x}| \geq u) \leq \sup_{u>0} u^p \sup_{\|x\|=1} \left\| \mathbb{E} \left( \mathbf{1}_{|D_{k,x}| \geq u} | \mathcal{F}_{k-1} \right) \right\|_\infty \leq c_2, \quad (3.7)$$

for any positive integer  $k$ . Taking  $\gamma = n^\alpha y$ ,  $u = n^\alpha y/r$  with  $r \in (0, \infty)$ , and  $v = 2nc_1$  in (3.3), we get

$$\mathbb{P} \left( \max_{1 \leq k \leq n} |M_{k,x}| \geq n^\alpha y \right) \leq c_3 \left( \frac{1}{y^p n^{\alpha p - 1}} + \frac{1}{y^{2r} n^{(2\alpha - 1)r}} \right), \quad (3.8)$$

for some positive constant  $c_3$ . Selecting  $r > (\alpha p - 1)/(2\alpha - 1)$ , the upper bound (3.2) follows directly from (3.8).

*The case  $p \in (1, 2)$ .* Let  $x \in S^{d-1}$ . We have

$$\log \|A_n x\| - n \lambda_\mu = \sum_{k=1}^n (X_{k,x} - \lambda_\mu) = \sum_{k=1}^n (D_{k,x} + R_{k,x}) := M_{n,x} + U_{n,x},$$



where

$$D_{n,x} = \sigma(Y_n, A_{n-1}x) - \int_G \sigma(g, A_{n-1}x) \mu(dg), \text{ and } R_{n,x} = \int_G \sigma(g, A_{n-1}x) \mu(dg) - \lambda_\mu.$$

Notice that  $\mathbb{E}(D_{k,x} | \mathcal{F}_{k-1}) = 0$ . We use now the basic inequality

$$\mathbb{P} \left( \max_{1 \leq k \leq n} |\log \|A_k x\| - k\lambda_\mu| \geq n^\alpha y \right) \leq \mathbb{P} \left( \max_{1 \leq k \leq n} |M_{k,x}| \geq n^\alpha y/2 \right) + \mathbb{P} \left( \max_{1 \leq k \leq n} |U_{k,x}| \geq n^\alpha y/2 \right). \quad (3.9)$$

We first deal with the second term on the right-hand side of (3.9). We shall need the following extension of Theorem 3 in [24]. The proof is given in Appendix.

**Theorem 3.2.** *Let  $p \in ]1, 2[$  and  $(X_k)_{k \in \mathbb{Z}}$  be a sequence of real-valued random variables in  $\mathbb{L}^p$  and adapted to a non-decreasing filtration  $(\mathcal{F}_k)_{k \in \mathbb{Z}}$ . Let  $S_i = X_1 + \dots + X_i$  and  $S_n^* = \max_{1 \leq i \leq n} |S_i|$ . Then, for any  $n \geq 1$ ,*

$$\|S_n^*\|_p \leq (2c_p + 1) \left( \sum_{j=1}^n \|X_j\|_p^p \right)^{1/p} + 2^{(p-1)/p} (2c_p + 1) \sum_{j=0}^{r-1} \left( \sum_{k=1}^{2^{r-j}} \|\mathbb{E}(S_{k2^j} - S_{(k-1)2^j} | \mathcal{F}_{(k-1)2^j})\|_p^p \right)^{1/p}, \quad (3.10)$$

where  $c_p = 2^{1/p} \frac{p}{p-1}$  and  $r$  is the unique positive integer such that  $2^{r-1} \leq n < 2^r$ .

For  $k \leq 0$ , set  $R_{k,x} = R_{0,x}$  and  $\mathcal{F}_k = \mathcal{F}_0$ . Observe that  $|R_{k,x}| \leq \int_G \log N(g) \mu(dg) < \infty$  for every  $k \geq 0$ . Hence we may apply Theorem 3.2 with  $X_k := R_{k,x}$ . With that choice, we have

$$S_{k2^j} - S_{(k-1)2^j} = \sum_{\ell=1}^{2^j} R_{(k-1)2^j + \ell, x},$$

and, using independence (twice),

$$|\mathbb{E}(R_{(k-1)2^j + \ell, x} | \mathcal{F}_{(k-1)2^j})| \leq \sup_{y \in S^{d-1}} \left| \int_G (\mathbb{E}(\sigma(g, A_{\ell-1}y)) - \lambda_\mu) \mu(dg) \right| = \sup_{y \in S^{d-1}} |\mathbb{E}(X_{\ell,y}) - \lambda_\mu|.$$

Let  $n \geq 1$  and  $r \geq 1$  be such that  $2^{r-1} \leq n < 2^r$ . We infer that there exists  $C_p > 0$ , such that

$$\begin{aligned} \left\| \max_{1 \leq k \leq n} |U_{k,x}| \right\|_p &\leq C_p n^{1/p} + C_p \sum_{j=0}^{r-1} 2^{(r-j)/p} \sum_{\ell=1}^{2^j} \sup_{\|y\|=1} |\mathbb{E}(X_{\ell,y}) - \lambda_\mu| \\ &\leq C_p n^{1/p} + \frac{C_p 2^{1/p}}{2^{1/p} - 1} n^{1/p} \sum_{\ell \geq 1} \frac{\sup_{\|y\|=1} |\mathbb{E}(X_{\ell,y}) - \lambda_\mu|}{\ell^{1/p}}. \end{aligned} \quad (3.11)$$

Recall that (3.1) holds for  $p \in (1, 2)$ . Hence, for any  $r < p$ , by (6) of [6],

$$\sum_{n \geq 1} n^{r-2} \sup_{\|y\|=1} |\mathbb{E}(X_{n,y}) - \lambda_\mu| < \infty. \quad (3.12)$$

Since  $p + 1/p > 2$  one can choose  $r$  close enough to  $p$  in such a way that  $-1/p < r - 2$ . In particular, it follows that

$$\sum_{\ell \geq 1} \ell^{-1/p} \sup_{\|y\|=1} |\mathbb{E}(X_{\ell,y}) - \lambda_\mu| < \infty. \quad (3.13)$$

Hence, using (3.11),

$$\left\| \max_{1 \leq i \leq n} |U_{i,x}| \right\|_p \leq \tilde{C}_p n^{1/p},$$

and

$$\sup_{\|x\|=1} \mathbb{P} \left( \max_{1 \leq k \leq n} |U_{k,x}| \geq y n^\alpha / 2 \right) \leq \frac{(2\tilde{C}_p)^p n}{y^p n^{p\alpha}}, \quad (3.14)$$

which ends the control of the second term on the right-hand side of (3.9).

We now deal with the first term on the right-hand side of (3.9), that is the martingale term. We shall need the following result (to be proved in Appendix). It is a maximal-version of Theorem 2.5 in [15] (a von Bahr-Esseen inequality for martingales having weak moments of order  $p \in (1, 2)$ ). For a real-valued random variable  $X$ , let  $\|X\|_{p,\infty} = \sup_{t>0} t(\mathbb{P}(|X| > t))^{1/p}$ .

**Proposition 3.3.** *Let  $(D_n)_{n \in \mathbb{N}}$  be a sequence of  $(\mathcal{F}_n)_{n \in \mathbb{N}}$ -martingale differences in weak- $L^p$ ,  $p \in (1, 2)$ . Then*

$$\mathbb{P} \left( \max_{1 \leq k \leq n} \left| \sum_{j=1}^k D_j \right| \geq y \right) \leq \frac{K}{y^p} \sum_{k=1}^n \|D_k\|_{p,\infty}^p,$$

where  $K = 4p/(p-1) + 8/(2-p)$ .

Now, since (3.1) holds, then so does (3.7). It follows from Proposition 3.3 that

$$\sup_{\|x\|=1} \mathbb{P} \left( \max_{1 \leq k \leq 2^n} |M_{k,x}| \geq n^\alpha y / 2 \right) \leq \frac{C}{y^p n^{\alpha p - 1}}, \quad (3.15)$$

for some positive constant  $C$ . The upper bound (3.2) follows from (3.9), (3.14) and (3.15).

*The case  $p = 2$ .* We start from (3.9). Note first that the upper bound (3.14) still holds for  $p = 2$ , with the same proof. We now deal with the first term on the right-hand side of (3.9). Instead of Proposition (3.3), we shall use the following result of Hao and Liu [17] (see also Theorem 14 in [6]): if  $\mathbb{P}(|D_{k,x}| > y) \leq \mathbb{P}(X > y)$  for any  $y > 0$  and some positive random variable  $X$ , then, for every  $q > 1$ , every  $\gamma \in (1, 2]$  and every  $L \in \mathbb{N}$ , there exists  $C > 0$ , such that for every  $n \geq 1$  and every  $\lambda > 0$ ,

$$\begin{aligned} \mathbb{P} \left( \max_{1 \leq k \leq n} |M_{k,x}| \geq \lambda \right) &\leq n \mathbb{P} \left( X > \frac{\lambda}{4(L+1)} \right) \\ &+ \frac{C}{(\lambda)^{q\gamma(L+1)/(q+L)}} \left\| \mathbb{E}(|D_{1,x}|^\gamma | \mathcal{F}_0) + \cdots + \mathbb{E}(|D_{n,x}|^\gamma | \mathcal{F}_{n-1}) \right\|_q^{q(L+1)/(q+L)}. \end{aligned} \quad (3.16)$$

We apply (3.16) with  $X = \log N(Y_1) + \mathbb{E}(\log N(Y_1))$ . Since (3.1) holds with  $p = 2$ , then  $X$  has a weak moment of order 2, and, for every  $\gamma \in (1, 2)$ , there exists  $C_\gamma > 0$  such that for every  $n \geq 1$ ,

$$\|\mathbb{E}(|D_{n,x}|^\gamma | \mathcal{F}_{n-1})\|_\infty \leq C_\gamma.$$

Hence, for every integer  $L$  and every  $q > 1$ , there exists  $C > 0$  such that

$$\mathbb{P}\left(\max_{1 \leq k \leq n} |M_{k,x}| \geq n^\alpha y/2\right) \leq n\mathbb{P}\left(X \geq \frac{n^\alpha y}{8(L+1)}\right) + \frac{Cn^{q(L+1)/(q+L)}}{(n^\alpha y)^{q\gamma(L+1)/(q+L)}}.$$

Since  $\alpha > 1/2$  one may find  $\gamma \in (1, 2)$ , such that  $\gamma\alpha > 1$ . For such a choice, taking  $q = L$  large enough, we obtain the desired result.  $\diamond$

## 4 The case of strong moments of order $p \geq 1$

In this section, we prove some results in the spirit of Baum and Katz [2] for the quantity (1.5).

**Theorem 4.1.** *Let  $p \geq 1$  and assume that*

$$\int (\log N(g))^p \mu(dg) < \infty. \quad (4.1)$$

*Let  $\alpha \in (1/2, 1]$  and  $\alpha \geq 1/p$ . Then for any  $y > 0$*

$$\sum_{n \geq 1} n^{\alpha p - 2} \sup_{\|x\|=1} \mathbb{P}\left(\max_{1 \leq k \leq n} |\log \|A_k x\| - k\lambda_\mu| > n^\alpha y\right) < \infty. \quad (4.2)$$

**Remark 4.1.** Theorem 4.1 is due to Benoist-Quint [4] in the case where  $\alpha = 1$  and  $p > 1$ .

**Remark 4.2.** Let us recall a well known consequence of (4.2), when  $p \in [1, 2)$  and  $\alpha = 1/p$ . The sequence  $\max_{1 \leq k \leq n} |\log \|A_k x\| - k\lambda_\mu|$  being non-decreasing, Inequality (4.2) with  $\alpha = 1/p$  is equivalent to

$$\sum_{N \geq 1} \sup_{\|x\|=1} \mathbb{P}\left(\max_{1 \leq k \leq 2^N} |\log \|A_k x\| - k\lambda_\mu| > 2^{N/p} y\right) < \infty. \quad (4.3)$$

This implies that, for any  $x \in S^{d-1}$ , the sequence  $(2^{-N/p} \max_{1 \leq k \leq 2^N} |\log \|A_k x\| - k\lambda_\mu|)_{N \geq 1}$  converges completely. It follows easily that, for any  $x \in S^{d-1}$ ,  $n^{-1/p}(\log \|A_k x\| - k\lambda_\mu)$  converges to 0 almost surely as  $n \rightarrow \infty$ . Hence (4.2) is a more precise statement than Theorem 1(i) of [6].

Of course, (4.2) does not hold for  $p = 2$  and  $\alpha = 1/2$ . Instead, we have the following result, which implies a bounded law of the iterated logarithm.

**Theorem 4.2.** *Assume that  $\int (\log N(g))^2 \mu(dg) < \infty$ , and let  $V$  be defined as in Theorem 2.3. Then for any  $y > \sqrt{V} \geq 0$ , we have*

$$\sum_{n \geq 1} \frac{1}{n} \sup_{\|x\|=1} \mathbb{P}\left(\max_{1 \leq k \leq n} |\log \|A_k x\| - k\lambda_\mu| > y\sqrt{2n \log \log n}\right) < \infty. \quad (4.4)$$

**Remark 4.3.** From (4.4) one can easily infer that, for any  $x \in S^{d-1}$ ,

$$\limsup_{n \rightarrow \infty} \frac{|\log \|A_n x\| - n\lambda_\mu|}{\sqrt{2n \log \log n}} \leq \sqrt{V}, \quad \text{almost surely.}$$

Of course, this is a less precise result than the compact law of the iterated logarithm, which also holds provided  $\int (\log N(g))^2 \mu(dg) < \infty$  (for instance, this is a consequence of Theorem 1(iii), case  $p = 2$ , of [6]). Note however that (4.4) and the compact law of the iterated logarithm are two different results, which cannot be deduced from one another.

**Proof of Theorem 4.1.** The proof follows the line of that of Theorem 3.1.

*The case  $p \geq 2$ .* In that case the decomposition (2.6) holds, and it is equivalent to prove (4.2) for  $M_{k,x}$  instead of  $(\log \|A_k x\| - k\lambda_\mu)$ .

Starting from (2.5) and (3.5), we see that

$$\sup_{\|x\|=1, k \geq 1} \mathbb{P}(|D_{k,x}| \geq n^\alpha y/r) \leq \mu\{\log N \geq n^\alpha y/2r\} + \mathbf{1}_{n^\alpha y \leq 2r(2\|\psi\|_\infty + |\lambda_\mu|)}. \quad (4.5)$$

Taking  $\gamma = n^\alpha y$ ,  $u = n^\alpha y/r$  with  $r > 0$ , and  $v = 2nc_1$  (cf. (3.6) for the definition of  $c_1$ ) in (3.3), we get

$$\mathbb{P}\left(\max_{1 \leq k \leq n} |M_{k,x}| \geq n^\alpha y\right) \leq n\mu\{\log N \geq n^\alpha y/2r\} + n\mathbf{1}_{n^\alpha y \leq 2r(2\|\psi\|_\infty + |\lambda_\mu|)} + \frac{\kappa_1}{y^{2r} n^{(2\alpha-1)r}}, \quad (4.6)$$

for some  $\kappa_1 > 0$ . Interverting the sum and the integral, we see that

$$\sum_{n>0} n^{\alpha p-1} \mu\{\log N \geq n^\alpha y/2r\} \leq \frac{\kappa_2}{y^p} \int (\log N(g))^p \mu(dg), \quad (4.7)$$

for some positive constant  $\kappa_2$  depending only on  $r$ . Taking  $r > (\alpha p - 1)/(2\alpha - 1)$  in (4.6) and using the upper bound (4.7), the proof of (4.2) is complete for  $p \geq 2$ .

*The case  $p \in (1, 2)$ .* We start again from (3.9). If (4.1) holds for  $p \in (1, 2)$  then, by (6) of [6], (3.12) holds with  $r = p$ . Since  $p + 1/p > 2$  and  $p < 2$ , there exists  $q$  such that  $p < q < 2$  and  $p + 1/q > 2$ . Hence, we infer that

$$\sum_{n \geq 1} n^{-1/q} \sup_{\|x\|=1} |\mathbb{E}(X_n) - \lambda_\mu| < \infty, \quad (4.8)$$

and, using (3.11) in  $\mathbb{L}^q$  rather than in  $\mathbb{L}^p$ , we infer that for every  $n \geq 1$

$$\left\| \max_{1 \leq i \leq n} |U_{i,x}| \right\|_q \leq C_q n^{1/q},$$

for some  $C_q > 0$ . Hence,

$$\sup_{\|x\|=1} \mathbb{P}\left(\max_{1 \leq k \leq n} |U_{k,x}| \geq y n^\alpha / 2\right) \leq \frac{(2C_q)^q n}{y^q n^{q\alpha}},$$

and, since  $q > p$ ,

$$\sum_{n \geq 1} n^{\alpha p-2} \sup_{\|x\|=1} \mathbb{P}\left(\max_{1 \leq k \leq n} |U_{k,x}| > n^\alpha y/2\right) < \infty. \quad (4.9)$$

It remains to deal with the first term on the right-hand side of (3.9). Applying (2.5) and (3.5) again, we see that  $|D_{k,x}|$  is uniformly (with respect to  $k$  and  $x$ ) stochastically bounded by the random variable  $\log N(Y_1) + |\lambda_\mu| + 2\|\psi\|_\infty$ . Following exactly the proof of Theorem 2 of [8], it follows that

$$\sum_{n \geq 1} n^{\alpha p - 2} \sup_{\|x\|=1} \mathbb{P} \left( \max_{1 \leq k \leq n} |M_{k,x}| > n^\alpha y / 2 \right) < \infty, \quad (4.10)$$

and the proof is complete for  $p \in (1, 2)$ .

*The case  $p = 1$ .* In that case  $\alpha = 1$  also. So, let us start from the inequality (3.9) with  $\alpha = 1$ .

The second term on the right-hand side of (3.9) may be handled thanks to a “maximal version” of Proposition 3.1 of [4], which implies that: for any  $y > 0$ , there exists constants  $A > 0$ ,  $\alpha > 0$  such that

$$\sup_{\|x\|=1} \mathbb{P} \left( \max_{1 \leq k \leq n} |U_{k,x}| \geq ny/2 \right) \leq Ae^{-\alpha n}. \quad (4.11)$$

Note that a direct application of Proposition 3.1 of [4] gives us (4.11) without the maximum over  $k$ . To prove (4.11), it is convenient to work on the projective space  $X := \mathbb{P}_{d-1}(\mathbb{R})$  rather than on  $S^{d-1}$ . Denote by  $\bar{x}$  the class (in  $X$ ) of  $x \in \mathbb{R}^d - \{0\}$ . Then, we define  $U_{n,\bar{x}} := U_{n,x}$ . Recall that, by our assumptions, the Markov chain  $(\overline{A_{n-1} \cdot x})_{n \geq 1}$  with (compact) state space  $X$  and transition probability given by  $Pf(\bar{x}) := \int_G f(\overline{g \cdot x}) \mu(dg)$  has a unique invariant probability  $\nu$ . In particular, for every continuous function  $f$  on  $X$ , the sequence

$$\left( \frac{1}{n} \mathbb{E} \left( \sum_{k=1}^n f(\overline{A_{k-1} \cdot x}) \right) \right)_{n \geq 1}$$

converges uniformly (with respect to  $\bar{x}$ ) to  $\nu(f)$ . Hence, there exists an integer  $m \geq 1$  such that

$$\sup_{\bar{x} \in X} |\mathbb{E}(U_{m,\bar{x}})| < my/4. \quad (4.12)$$

The maximal inequality (4.11) follows then by applying Lemma 23 in [9].

Let us deal with the first term on the right-hand side of (3.9). Let  $\Gamma_n := \cup_{k=1}^n \{\log(N(Y_k)) \geq yn\}$  and notice that on  $\Gamma_n^c$ ,

$$\sigma(Y_k, A_{k-1}x) = \sigma(Y_k, A_{k-1}x) \mathbf{1}_{\{\log(N(Y_k)) < yn\}}.$$

Define

$$\tilde{M}_{k,x} := \sum_{j=1}^k \left( \sigma(Y_j, A_{j-1}x) \mathbf{1}_{\{\log(N(Y_j)) < yn\}} - \int_G \sigma(g, A_{j-1}x) \mathbf{1}_{\{\log(N(g)) < yn\}} \mu(dg) \right),$$

and note that  $(\tilde{M}_{k,x})_{1 \leq k \leq n}$  is a martingale. Let

$$I(n) = \int_G \log(N(g)) \mathbf{1}_{\{\log(N(g)) \geq yn\}} \mu(dg),$$

and note that

$$\mathbb{P} \left( \max_{1 \leq k \leq n} |M_{k,x}| \geq ny/2 \right) \leq \mathbb{P}(\Gamma_n) + \mathbb{P} \left( \max_{1 \leq k \leq n} |\tilde{M}_{k,x}| \geq ny/4 \right) + \mathbf{1}_{\{I(n) \geq y/4\}}.$$

Using Doob's maximal inequality, we infer that

$$\mathbb{P} \left( \max_{1 \leq k \leq n} |M_{k,x}| \geq ny/2 \right) \leq \mathbb{P}(\Gamma_n) + \frac{64}{n^2 y^2} \mathbb{E} \left( \tilde{M}_{n,x}^2 \right) + \mathbf{1}_{\{I(n) \geq y/4\}}.$$

The last term on the right-hand side is equal to 0 for  $n$  large enough since (4.1) holds with  $p = 1$ .

Now,  $\mathbb{P}(\Gamma_n) \leq n\mathbb{P}(\log N(Y_1) \geq yn)$ . Hence it is standard that  $\sum_{n \geq 1} n^{-1}\mathbb{P}(\Gamma_n) < \infty$ , since (4.1) holds with  $p = 1$ .

On the other hand,

$$\sup_{\|x\|=1} \mathbb{E} \left( \tilde{M}_{n,x}^2 \right) \leq n \int_G (\log N(g))^2 \mathbf{1}_{\{\log(N(g)) < yn\}} \mu(dg).$$

Then, it is also standard that  $\sum_{n \geq 1} n^{-3} \sup_{\|x\|=1} \mathbb{E}(\tilde{M}_{n,x}^2) < \infty$ , since (4.1) holds with  $p = 1$ .  $\diamond$

**Proof of Theorem 4.2.** In that case the decomposition (2.6) holds, and it is equivalent to prove (4.2) for  $M_{k,x}$  instead of  $(\log \|A_k x\| - k\lambda_\mu)$ . Moreover, we have  $V = \mathbb{E}(D_{1,x}^2)$ .

We shall proceed by truncation. Let  $y > \sqrt{V}$  and set  $\varepsilon := y - \sqrt{V}$ . Let  $n \geq 1$ . Let  $\alpha > 0$  be fixed for the moment. For every  $1 \leq k \leq n$ , define

$$\tilde{D}_{k,n,x} := D_{k,x} \mathbf{1}_{\{|D_{k,x}| \leq \alpha\sqrt{n}/\sqrt{\log \log n}\}} - \mathbb{E} \left( D_{k,x} \mathbf{1}_{\{|D_{k,x}| \leq \alpha\sqrt{n}/\sqrt{\log \log n}\}} | \mathcal{F}_{k-1} \right)$$

and

$$\tilde{M}_{k,n,x} = \sum_{j=1}^k \tilde{D}_{j,n,x}.$$

Then, using Markov's inequality and stationarity

$$\begin{aligned} \mathbb{P} \left( \max_{1 \leq k \leq n} |M_{k,x}| > y\sqrt{2n \log \log n} \right) &\leq \mathbb{P} \left( \max_{1 \leq k \leq n} |\tilde{M}_{k,n,x}| > (y - \varepsilon/2)\sqrt{2n \log \log n} \right) \\ &\quad + \mathbb{P} \left( \max_{1 \leq k \leq n} |M_{k,x} - \tilde{M}_{k,n,x}| > y\sqrt{2n \log \log n} \right) \\ &\leq \mathbb{P} \left( \max_{1 \leq k \leq n} |\tilde{M}_{k,n,x}| > (\varepsilon/2)\sqrt{2n \log \log n} \right) \\ &\quad + \frac{2n}{\varepsilon\sqrt{2n \log \log n}} \mathbb{E} \left( 2|D_{1,x}| \mathbf{1}_{\{|D_{1,x}| > \alpha\sqrt{n}/\sqrt{\log \log n}\}} \right). \end{aligned}$$

Now, starting from (2.5) and arguing as in (3.5),

$$\begin{aligned} \sup_{\|x\|=1, k \geq 1} \mathbb{E} \left( |D_{1,x}| \mathbf{1}_{\{|D_{1,x}| > \alpha\sqrt{n}/\sqrt{\log \log n}\}} \right) \\ \leq \mathbb{E} \left( (N(Y_1) + |\lambda_\mu| + 2\|\psi\|_\infty) \mathbf{1}_{\{N(Y_1) + |\lambda_\mu| + 2\|\psi\|_\infty > \alpha\sqrt{n}/\sqrt{\log \log n}\}} \right). \end{aligned}$$

Since  $\int (\log N(g))^2 \mu(dg) < \infty$ , it is now standard that

$$\sum_{n \geq 1} \frac{4}{\varepsilon \sqrt{2n \log \log n}} \sup_{\|x\|=1} \mathbb{E} \left( |D_{1,x}| \mathbf{1}_{\{|D_{1,x}| > \alpha \sqrt{n}/\sqrt{\log \log n}\}} \right) < \infty.$$

Hence, it remains to prove that

$$\sum_{n \geq 1} \frac{1}{n} \sup_{\|x\|=1} \mathbb{P} \left( \max_{1 \leq k \leq n} |\tilde{M}_{k,n,x}| > (y - \varepsilon/2) \sqrt{2n \log \log n} \right) < \infty. \quad (4.13)$$

We shall use the following sharper version of Haeusler's bound (3.3) (see the end of the Proof of Lemma 1 in [16]).

$$\begin{aligned} \mathbb{P} \left( \max_{1 \leq k \leq n} |\tilde{M}_{k,n,x}| \geq \gamma \right) &\leq \sum_{i=1}^n \mathbb{P} \left( |\tilde{D}_{i,n,x}| \geq u \right) + 2\mathbb{P} \left( \sum_{i=1}^n \mathbb{E}(\tilde{D}_{i,n,x}^2 | \mathcal{F}_{i-1}) \geq v \right) \\ &\quad + 2 \exp \left( \gamma u^{-1} - (\gamma u^{-1} + v u^{-2}) \log(\gamma u v^{-1} + 1) \right). \end{aligned} \quad (4.14)$$

Notice that for every  $1 \leq k \leq n$ ,  $|\tilde{D}_{k,n,x}| \leq 2\alpha \sqrt{n}/\sqrt{\log \log n}$  and that

$$\sum_{i=1}^n \mathbb{E}(\tilde{D}_{i,n,x}^2 | \mathcal{F}_{i-1}) \leq \sum_{i=1}^n \mathbb{E}(D_{i,x}^2 | \mathcal{F}_{i-1}). \quad (4.15)$$

By Proposition 3.1 of [4],

$$\sum_{n \geq 1} \sup_{\|x\|=1} \mathbb{P} \left( \sum_{i=1}^n \mathbb{E}(\tilde{D}_{i,n,x}^2 | \mathcal{F}_{i-1}) \geq n(\sqrt{V} + \varepsilon)^2 \right) < \infty. \quad (4.16)$$

We shall apply (4.14) with  $\gamma := (y - \varepsilon/2) \sqrt{2n \log \log n}$ ,  $v := (\sqrt{V} + \varepsilon)^2 n$  and  $u := 4\alpha \sqrt{n}/\sqrt{\log \log n}$ . Using that for every  $t \geq 0$ ,  $\log(1+t) \geq t - t^2/2$ , we infer that

$$\gamma u^{-1} - (\gamma u^{-1} + v u^{-2}) \log(\gamma u v^{-1} + 1) \leq -\frac{\gamma^2 v^{-1}}{2} (1 - \gamma u v^{-1}).$$

Since

$$\frac{\gamma^2 v^{-1}}{2 \log \log n} = \frac{(\sqrt{V} + \varepsilon/2)^2}{(\sqrt{V} + \varepsilon/4)^2} > 1,$$

and since  $\gamma u v^{-1} = 4\sqrt{2}\alpha(\sqrt{V} + \varepsilon/2)(\sqrt{V} + \varepsilon/4)^2 \rightarrow 0$  as  $\alpha \rightarrow 0$ , we can choose  $\alpha$  small enough in such a way that there exists  $\delta > 1$  for which

$$\exp \left( \gamma u^{-1} - (\gamma u^{-1} + v u^{-2}) \log(\gamma u v^{-1} + 1) \right) \leq (\log n)^{-\delta}. \quad (4.17)$$

Combining (4.14), (4.15), (4.16) and (4.17) we conclude that (4.13) holds.  $\diamond$

## 5 Appendix

### 5.1 Proof of Proposition 2.4

As we shall see the proposition is a consequence of the following more general result concerning the functional moderate deviation principle of an array of martingale differences.

**Theorem 5.1.** *Let  $(d_{i,n})_{1 \leq i \leq n}$  be an array of square-integrable martingale differences, adapted to an array of filtrations  $(\mathcal{F}_{i,n})_{0 \leq i \leq n}$ . Let  $(b_n)_{n \geq 0}$  be a sequence of positive numbers such that  $b_n/\sqrt{n} \rightarrow \infty$  and  $b_n/n \rightarrow 0$  as  $n \rightarrow \infty$ , and let*

$$\bar{Z}_n = \left\{ \frac{d_{1,n} + \cdots + d_{[nt],n}}{b_n}, t \in [0, 1] \right\}.$$

*Suppose the conditions (2.16) and (2.17) satisfied. In addition, assume that*

$$\limsup_{n \rightarrow \infty} \frac{n}{b_n^2} \log \mathbb{P} \left( \max_{1 \leq k \leq n} |d_{k,n}| > b_n \right) = -\infty. \quad (5.1)$$

*and that, for any  $\lambda > 0, \delta > 0$ ,*

$$\limsup_{n \rightarrow \infty} \frac{n}{b_n^2} \log \mathbb{P} \left( \frac{n}{b_n^2} \sum_{k=1}^n \mathbb{E} \left( e^{\frac{\lambda b_n |d_{k,n}|}{n}} \mathbf{1}_{nb_n^{-1} < |d_{k,n}| \leq b_n} | \mathcal{F}_{k-1,n} \right) > \delta \right) = -\infty. \quad (5.2)$$

*Then, the functional moderate deviation principle (2.19) holds.*

**Proof of Theorem 5.1.** The proof will be done with the help of a truncature argument, using Puhalskii's functional moderate deviation principle for the main part and proving that the other parts have negligible contributions.

First, to soothe the notations, we suppress the index  $n$  and we denote  $d_k = d_{k,n}$  and  $\mathcal{F}_k = \mathcal{F}_{k,n}$ . We use a tuncation of the variables  $d_k$  as follows: for all  $1 \leq k \leq n$ , let

$$\bar{d}_k := d_k \mathbf{1}_{|d_k| \leq nb_n^{-1}} - \mathbb{E} \left( d_k \mathbf{1}_{|d_k| \leq nb_n^{-1}} | \mathcal{F}_{k-1} \right),$$

$$d'_k := d_k \mathbf{1}_{nb_n^{-1} < |d_k| \leq b_n} - \mathbb{E} \left( d_k \mathbf{1}_{nb_n^{-1} < |d_k| \leq b_n} | \mathcal{F}_{k-1} \right)$$

and

$$d''_k := d_k \mathbf{1}_{|d_k| > b_n} - \mathbb{E} \left( d_k \mathbf{1}_{|d_k| > b_n} | \mathcal{F}_{k-1} \right).$$

With these notations, we clearly have that, for any  $t \in [0, 1]$

$$Z_n(t) = b_n^{-1} \sum_{k=1}^{[nt]} d_k = \bar{Z}_n(t) + Z'_n(t) + Z''_n(t),$$

with  $\bar{Z}_n(t) = b_n^{-1} \sum_{k=1}^{[nt]} \bar{d}_k$ ,  $Z'_n(t) = b_n^{-1} \sum_{k=1}^{[nt]} d'_k$  and  $Z''_n(t) = b_n^{-1} \sum_{k=1}^{[nt]} d''_k$ . Notice first that

$$b_n^{-1} \sum_{j=1}^n \mathbb{E} (|d_k| \mathbf{1}_{|d_k| > b_n} | \mathcal{F}_{k-1}) \leq \frac{n}{b_n^2} \frac{1}{n} \sum_{j=1}^n \mathbb{E} (d_k^2 \mathbf{1}_{|d_k| > b_n} | \mathcal{F}_{k-1})$$



and that for any  $\delta > 0$ ,

$$\frac{n}{b_n^2} \log \mathbb{P} \left( b_n^{-1} \sum_{k=1}^n |d_k| \mathbf{1}_{|d_k| > b_n} \geq \delta \right) \leq \frac{n}{b_n^2} \log \mathbb{P} \left( \max_{1 \leq k \leq n} |d_k| > b_n \right).$$

Hence, by taking into account conditions (2.17) and (5.1), we can deduce that the process  $Z_n''$  has a negligible contribution to the functional moderate deviation principle (see Theorem 4.2.13 in [10]).

On another hand, note that  $(\bar{d}_k)_{1 \leq k \leq n}$  is a triangular sequence of martingale differences such that  $\|\bar{d}_k\|_\infty \leq 2n/b_n$ . Using conditions (2.16) and (2.17), we can apply the functional moderate deviation principle of Puhalskii [22] which entails that  $\bar{Z}_n$  satisfies (2.19). Therefore to end the proof it remains to show that the process  $Z_n'$  has a negligible contribution to the functional moderate deviation principle; that is: for any  $\delta > 0$ ,

$$\limsup_{n \rightarrow \infty} \frac{n}{b_n^2} \log \mathbb{P} \left( \sup_{t \in [0,1]} |Z_n'(t)| > \delta \right) = -\infty. \quad (5.3)$$

Observe that

$$b_n^{-1} \sum_{j=1}^n \mathbb{E} \left( |d_k| \mathbf{1}_{nb_n^{-1} < |d_k| \leq b_n} | \mathcal{F}_{k-1} \right) \leq \frac{1}{n} \sum_{j=1}^n \mathbb{E} \left( d_k^2 \mathbf{1}_{|d_k| > nb_n^{-1}} | \mathcal{F}_{k-1} \right),$$

which by using condition (2.17) implies that (5.3) will hold if we can prove that, for any  $\delta > 0$ ,

$$\limsup_{n \rightarrow \infty} \frac{n}{b_n^2} \log \mathbb{P} \left( b_n^{-1} \sum_{k=1}^n |d_k| \mathbf{1}_{nb_n^{-1} < |d_k| \leq b_n} > \delta \right) = -\infty. \quad (5.4)$$

With this aim, we use the arguments developed in the proof of Proposition 1 in [23]. For the sake of clarity, let us give some details. Take  $\lambda$  a positive number and set  $Y_{k,\lambda} := \frac{2\lambda b_n}{n} |d_k| \mathbf{1}_{nb_n^{-1} < |d_k| \leq b_n}$ . We have

$$\begin{aligned} \mathbb{P} \left( b_n^{-1} \sum_{k=1}^n |d_k| \mathbf{1}_{nb_n^{-1} < |d_k| \leq b_n} > \delta \right) &= \mathbb{P} \left( \sum_{k=1}^n Y_{k,\lambda} > \frac{2\lambda \delta b_n^2}{n} \right) \\ &\leq \mathbb{P} \left( \sum_{k=1}^n \{Y_{k,\lambda} - \log \mathbb{E}(e^{Y_{k,\lambda}} | \mathcal{F}_{k-1})\} > \frac{\lambda \delta b_n^2}{n} \right) + \mathbb{P} \left( \sum_{k=1}^n \log \mathbb{E}(e^{Y_{k,\lambda}} | \mathcal{F}_{k-1}) > \frac{\lambda \delta b_n^2}{n} \right). \end{aligned}$$

Since the  $Y_{k,\lambda}$  are  $\mathcal{F}_k$ -measurable, we have

$$\mathbb{E} \left( \frac{\prod_{k=1}^n e^{Y_{k,\lambda}}}{\prod_{k=1}^n \mathbb{E}(e^{Y_{k,\lambda}} | \mathcal{F}_{k-1})} \right) = 1.$$

Hence

$$\limsup_{n \rightarrow \infty} \frac{n}{b_n^2} \log \mathbb{P} \left( \sum_{k=1}^n \{Y_{k,\lambda} - \log \mathbb{E}(e^{Y_{k,\lambda}} | \mathcal{F}_{k-1})\} > \frac{\lambda \delta b_n^2}{n} \right) \leq -\lambda \delta,$$

which is going to  $-\infty$  by letting  $\lambda$  tend to  $\infty$ . Hence, to prove (5.4) (and then (5.3)), it suffices to show that, for any positive  $\lambda, \delta$ ,

$$\limsup_{n \rightarrow \infty} \frac{n}{b_n^2} \log \mathbb{P} \left( \sum_{k=1}^n \log \mathbb{E}(e^{Y_{k,\lambda}} | \mathcal{F}_{k-1}) > \frac{\lambda \delta b_n^2}{n} \right) = -\infty.$$

This holds under condition (5.2) by taking into account that  $e^{x\mathbf{1}_A} - 1 = (e^x - 1)\mathbf{1}_A$  and also that  $\log(1 + u) \leq u$  for any  $x > 0, u > 0$ . The proof of Theorem 5.1 is therefore complete.  $\diamond$

**End of the proof of Proposition 2.4.** We start with some observations. Obviously condition (5.2) holds under the stronger one: for any  $\lambda > 0$ ,

$$\limsup_{n \rightarrow \infty} \frac{n}{b_n^2} \sum_{k=1}^n \left\| \mathbb{E} \left( e^{\frac{\lambda b_n |d_{k,n}|}{n}} \mathbf{1}_{nb_n^{-1} < |d_{k,n}| \leq b_n} | \mathcal{F}_{k-1,n} \right) \right\|_{\infty} = 0.$$

Note now that this condition is equivalent to the following one. There is a constant  $C$  with the following property: for any  $\lambda > 0$  there exists a positive integer  $N(\lambda)$  such that for  $n > N(\lambda)$ ,

$$\frac{n}{b_n^2} \sum_{k=1}^n \left\| \mathbb{P}(|d_{k,n}| > unb_n^{-1} | \mathcal{F}_{k-1,n}) \right\|_{\infty} \leq C \exp(-\lambda u) \quad \text{for all } 1 \leq u \leq b_n^2/n. \quad (5.5)$$

(The proof of this equivalence can be done by following the proof of Comment 6 in [21]). To end the proof of the proposition, it remains to show that condition (2.18) implies (5.5) (since it obviously implies condition (5.1)). Under the regularity conditions (2.11), this can be achieved by following the lines of the proof of Corollary 7 in [21] (by taking  $s_n = \sqrt{n}$ ,  $k_n = n$  and  $a_n = n/b_n^2$ ).  $\diamond$

## 5.2 Proof of Theorem 3.2

**Proof.** The proof follows the lines of the proof of Theorem 3 in [24], but in the non-stationary setting, and is then done by induction. For  $n = 1$ , the inequality is clearly true. Assume that the inequality holds up to  $n - 1$  for any sequence  $(X_k)_{k \in \mathbb{Z}}$  of real-valued random variables in  $\mathbb{L}^p$  and adapted to a non-decreasing filtration  $(\mathcal{F}_k)_{k \in \mathbb{Z}}$ , and let us prove it for  $n$ . Set  $a_p = 2c_p + 1$ . By the triangle inequality

$$S_n^* \leq \max_{1 \leq k \leq n} \left| \sum_{i=1}^k [X_i - \mathbb{E}(X_i | \mathcal{F}_{i-1})] \right| + \max_{1 \leq k \leq n} \left| \sum_{i=1}^k \mathbb{E}(X_i | \mathcal{F}_{i-1}) \right|. \quad (5.6)$$

By von Bahr-Esseen's inequality together with Doob's maximal inequality for martingales,

$$\left\| \max_{1 \leq k \leq n} \left| \sum_{i=1}^k (X_i - \mathbb{E}(X_i | \mathcal{F}_{i-1})) \right| \right\|_p \leq c_p \left( \sum_{i=1}^n \|X_i - \mathbb{E}(X_i | \mathcal{F}_{i-1})\|_p^p \right)^{1/p} \leq 2c_p \left( \sum_{i=1}^n \|X_i\|_p^p \right)^{1/p}. \quad (5.7)$$

To estimate the impact of the second term in the right-hand side of (5.6), we start by writing  $n = 2m$ , or  $n = 2m + 1$  according to a value odd or even of  $n$ . Notice that

$$\left\| \max_{1 \leq k \leq n} \left| \sum_{j=1}^k \mathbb{E}(X_j | \mathcal{F}_{j-1}) \right| \right\|_p \leq \left\| \max_{1 \leq k \leq m} \left| \sum_{i=1}^{2k} \mathbb{E}(X_i | \mathcal{F}_{i-1}) \right| \right\|_p + \left\| \max_{0 \leq k \leq m} |\mathbb{E}(X_{2k+1} | \mathcal{F}_{2k})| \right\|_p. \quad (5.8)$$

The second term in the right hand side of (5.8) is estimated in a trivial way:

$$\left\| \max_{0 \leq k \leq m} |\mathbb{E}(X_{2k+1} | \mathcal{F}_{2k})| \right\|_p \leq \left( \sum_{k=0}^m \|\mathbb{E}(X_{2k+1} | \mathcal{F}_{2k})\|_p^p \right)^{1/p} \leq \left( \sum_{i=1}^n \|X_i\|_p^p \right)^{1/p}, \quad (5.9)$$

since  $m$  is such that  $n = 2m$  or  $n = 2m + 1$ . For the first term in the right hand side of (5.8), we set

$$Y_i = \mathbb{E}(X_{2i-1}|\mathcal{F}_{2i-2}) + \mathbb{E}(X_{2i}|\mathcal{F}_{2i-1}), \quad W_i = \sum_{j=1}^i Y_j \text{ and } \mathcal{G}_i = \mathcal{F}_{2i-1},$$

and we note that

$$\left\| \max_{1 \leq k \leq m} \left| \sum_{i=1}^{2k} \mathbb{E}(X_i|\mathcal{F}_{i-1}) \right| \right\|_p = \left\| \max_{1 \leq k \leq m} \left| \sum_{i=1}^k Y_i \right| \right\|_p.$$

In addition,  $(Y_k)_{k \in \mathbb{Z}}$  is a sequence of real-valued random variables in  $\mathbb{L}^p$  and adapted to the non-decreasing filtration  $(\mathcal{G}_k)_{k \in \mathbb{Z}}$ . By the induction hypothesis, noticing that  $m < 2^{r-1} \leq n$ ,

$$\left\| \max_{1 \leq k \leq m} \left| \sum_{i=1}^k Y_i \right| \right\|_p \leq a_p \left( \sum_{j=1}^m \|Y_j\|_p^p \right)^{1/p} + 2^{(p-1)/p} a_p \sum_{j=0}^{r-2} \left( \sum_{k=1}^{2^{r-1-j}} \|\mathbb{E}(W_{k2^j} - W_{(k-1)2^j}|\mathcal{G}_{(k-1)2^j})\|_p^p \right)^{1/p}.$$

But

$$\|\mathbb{E}(W_{k2^j} - W_{(k-1)2^j}|\mathcal{G}_{(k-1)2^j})\|_p \leq \|\mathbb{E}(S_{k2^{j+1}} - S_{(k-1)2^{j+1}}|\mathcal{F}_{(k-1)2^{j+1}})\|_p.$$

On another hand,

$$\sum_{j=1}^m \|Y_j\|_p^p \leq 2^{p-1} \sum_{i=1}^m (\|\mathbb{E}(X_{2i-1}|\mathcal{F}_{2i-2})\|_p^p + \|\mathbb{E}(X_{2i}|\mathcal{F}_{2i-1})\|_p^p) \leq 2^{p-1} \sum_{i=1}^n \|\mathbb{E}(X_i|\mathcal{F}_{i-1})\|_p^2.$$

Therefore

$$\begin{aligned} \left\| \max_{1 \leq k \leq m} \left| \sum_{i=1}^{2k} \mathbb{E}(X_i|\mathcal{F}_{i-1}) \right| \right\|_p &\leq 2^{(p-1)/p} a_p \left( \sum_{i=1}^n \|\mathbb{E}(X_i|\mathcal{F}_{i-1})\|_p^p \right)^{1/p} \\ &\quad + 2^{(p-1)/p} a_p \sum_{j=1}^{r-1} \left( \sum_{k=1}^{2^{r-j}} \|\mathbb{E}(S_{k2^j} - S_{(k-1)2^j}|\mathcal{F}_{(k-1)2^j})\|_p^2 \right)^{1/2}, \end{aligned}$$

which gives since  $n < 2^r$ ,

$$\left\| \max_{1 \leq k \leq m} \left| \sum_{i=1}^{2k} \mathbb{E}(X_i|\mathcal{F}_{i-1}) \right| \right\|_p \leq 2^{(p-1)/p} a_p \sum_{j=0}^{r-1} \left( \sum_{k=1}^{2^{r-j}} \|\mathbb{E}(S_{k2^j} - S_{(k-1)2^j}|\mathcal{F}_{(k-1)2^j})\|_p^2 \right)^{1/2}. \quad (5.10)$$

So, overall, starting from (5.6) and taking into account the upper bounds (5.7), (5.8), (5.9) and (5.10), Inequality (3.10) follows proving the induction hypothesis at step  $n$ .  $\diamond$

### 5.3 Proof of Proposition 3.3

Let  $M > 0$ . For every  $1 \leq k \leq n$ , define

$$\tilde{D}_k := D_k \mathbf{1}_{\{|D_k| \leq y\}} - \mathbb{E}(D_k \mathbf{1}_{\{|D_k| \leq y\}}|\mathcal{F}_{k-1}),$$

so that  $(\tilde{D}_k)_{1 \leq k \leq n}$  is a sequence of martingale differences. We have

$$\mathbb{P} \left( \max_{1 \leq k \leq n} \left| \sum_{j=1}^k (D_j - \tilde{D}_j) \right| \geq y/2 \right) \leq \frac{4}{y} \sum_{k=1}^n \mathbb{E} (|D_k| \mathbf{1}_{\{|D_k| > y\}}) .$$

Now,

$$\mathbb{E} (|D_k| \mathbf{1}_{\{|D_k| > y\}}) = \int_0^y \mathbb{P}(|D_k| > y) dt + \int_y^{+\infty} \mathbb{P}(|D_k| > t) dt \leq \frac{p}{p-1} \|D_k\|_{p,\infty}^p y^{1-p} .$$

Hence,

$$\mathbb{P} \left( \max_{1 \leq k \leq n} \left| \sum_{j=1}^k (D_j - \tilde{D}_j) \right| \geq y/2 \right) \leq \frac{4p}{y^p(p-1)} \sum_{k=1}^n \|D_k\|_{p,\infty}^p . \quad (5.11)$$

On another hand, by Doob's maximal inequality,

$$\mathbb{P} \left( \max_{1 \leq k \leq n} \left| \sum_{j=1}^k \tilde{D}_j \right| \geq y/2 \right) \leq \frac{4}{y^2} \sum_{k=1}^n \mathbb{E} (D_k^2 \mathbf{1}_{\{|D_k| \leq y\}}) .$$

Now,

$$\mathbb{E} (D_k^2 \mathbf{1}_{\{|D_k| \leq y\}}) \leq \int_0^y 2t \mathbb{P}(|D_k| > t) dt \leq \frac{2}{2-p} \|D_j\|_{p,\infty}^p y^{2-p} .$$

Hence,

$$\mathbb{P} \left( \max_{1 \leq k \leq n} \left| \sum_{j=1}^k \tilde{D}_j \right| \geq y/2 \right) \leq \frac{8}{y^p(2-p)} \sum_{k=1}^n \|D_k\|_{p,\infty}^p . \quad (5.12)$$

The result follows from (5.11) and (5.12).

## 6 General cocycles

It turns out that all the results obtained under moments greater than 2 made use of a martingale-coboundary decomposition with bounded (in  $\mathbb{L}^\infty$ ) coboundary and of the fact that we study partial sums associated with a cocycle. Another ingredient of general nature used in the proofs is Proposition 3.1 of Benoist-Quint [4]. In particular all the results obtained under moments greater than 2 may be generalized to cocycles admitting such a martingale-coboundary decomposition. Such cocycles are called *centerable* in [4].

We shall also give sufficient conditions under which the results under moments weaker than 2 holds for general cocycles.

Let us describe the situations that should be considered in the sequel.

Let  $G$  be a locally compact second countable group. Let  $X$  be compact and second countable. Assume that  $G$  acts continuously on  $X$  and denote that action by  $g \cdot x$ .

Let  $\sigma : G \times X \rightarrow \mathbb{R}$  be a cocycle, meaning that it satisfies the equality (1.4). We shall only be concerned with continuous cocycles. Given a continuous cocycle, define  $\sigma_{\sup}(g) := \sup_{u \in X} |\sigma(g, u)|$  for every  $g \in G$ .

Let  $\mu$  be a probability measure on the Borel sets of  $G$ .

Assume that there exists a *unique*  $\mu$ -invariant probability  $\nu$  on the Borel sets of  $X$ , that is a unique probability satisfying (1.1).

Let  $(\Omega, \mathcal{F}, \mathbb{P})$  be a probability space. Assume that there exists a sequence  $(Y_n)_{n \geq 1}$  of iid random variables on  $(\Omega, \mathcal{F}, \mathbb{P})$  taking values in  $G$  with common law  $\mu$ . Define  $A_n := Y_n \cdots Y_1$  for every  $n \geq 1$  and  $A_0 = \mathbf{e}$  the neutral element of  $G$ .

Our goal is to study the sequence defined by

$$S_{n,u} := \sigma(A_n, u) = \sum_{k=0}^{n-1} \sigma(Y_{k+1}, A_k \cdot u) \quad \forall n \geq 1, \forall u \in X.$$

**Definition 6.1.** We say that  $\sigma$  is *centerable* if  $\sigma_{\sup} \in \mathbb{L}^1(\mu)$  and if there exist a cocycle  $\sigma_0$  and a continuous function  $\psi$  on  $X$  such that  $\int_G \sigma_0(g, u) \mu(dg) = \lambda_\mu$  for every  $u \in X$ , where  $\lambda_\mu := \int_{G \times X} \sigma(h, v) \mu(dh) \nu(dv)$ , and

$$\sigma(g, u) = \sigma_0(g, u) + \psi(u) - \psi(g \cdot u) \quad \forall (g, u) \in G \times X. \quad (6.1)$$

**Remark 6.1.** A sufficient condition for  $\sigma$  to be centerable is Gordin's condition:

$$\sum_{n \geq 0} \sup_{u \in X} |\mathbb{E}(\sigma(Y_{n+1}, A_n \cdot u)) - \lambda_\mu| = \sum_{n \geq 0} \sup_{u \in X} \left| \int_{G \times G} \sigma(g, g' \cdot u) \mu(dg) \mu^{*n}(dg) - \lambda_\mu \right| < \infty.$$

- Assume that there exist  $r > 0$  and  $\delta > 0$  such that

$$\int_G e^{\delta \sigma_{\sup}^r(g)} \mu(dg) < \infty.$$

If  $\sigma$  is centerable, then the conclusions of Theorem 2.1 and Theorem 2.2 hold with  $S_{k,u}$  in place of  $\log \|A_k x\|$  for the corresponding value of  $r > 0$ .

- Assume  $\sigma_{\sup} \in \mathbb{L}^2(\mu)$  and

$$\limsup_{n \rightarrow \infty} \frac{n}{b_n^2} \log n \mu \{ \sigma_{\sup} > b_n \} = -\infty, \quad (6.2)$$

for some sequence  $(b_n)_{n \geq 1}$  satisfying (2.11). Then, if  $\sigma$  is centerable, the conclusion of Theorem 2.3 holds with  $S_{[nt],u}$  in place of  $\log \|A_{[nt]} x\|$ .

In the same way, if  $\sigma$  is centerable, in the case of weak moments of order  $p > 2$  (resp. strong moments of order  $p \geq 2$ ), the conclusion of Theorem 3.1 (resp. of Theorem 4.1) holds with  $S_{k,u}$  in place of  $\log \|A_k x\|$ .

Let us now mention results under weak moments of order  $p$ ,  $1 < p < 2$  or under (strong) moments of order  $1 \leq p < 2$ .

- Let  $1 < p < 2$ . Assume that

$$\sup_{t>0} t^p \mu\{\sigma_{\text{sup}} > t\} < \infty,$$

and that

$$\sum_{n \geq 1} n^{-1-1/p} \sup_{u \in X} |\mathbb{E}(\sigma(Y_n, A_{n-1} \cdot u)) - \lambda_\mu| < \infty.$$

Then the conclusion of Theorem 3.1 holds with  $S_{k,u}$  in place of  $\log \|A_k x\|$ .

- Let  $1 \leq p < 2$ . Assume that  $\sigma_{\text{sup}} \in \mathbb{L}^p(\mu)$  and that there exists  $q > p$  such that

$$\sum_{n \geq 1} n^{-1-1/q} \sup_{u \in X} |\mathbb{E}(\sigma(Y_n, A_{n-1} \cdot u)) - \lambda_\mu| < \infty.$$

Then the conclusion of Theorem 4.1 holds with  $S_{k,u}$  in place of  $\log \|A_k x\|$ .

## References

- [1] M. A. Arcones, *The large deviation principle for stochastic processes. I*, Teor. Veroyatnost. i Primenen 47 (2002) 727-746.
- [2] L. B. Baum and M. Katz, *Convergence rates in the law of large numbers*, Trans. Amer. Math. Soc. 120 (1965) 108-123.
- [3] Y. Benoist and J.-F. Quint, *Random walks on reductive groups*, manuscript, <http://www.math.u-psud.fr/~benoist/prepubli/prepublication.html>
- [4] Y. Benoist and J.-F. Quint, *Central limit theorem for linear groups*, Ann. Probab. 44 (2016) 1308-1340.
- [5] P. Bougerol and J. Lacroix, *Products of random matrices with applications to Schrödinger operators*. Progress in Probability and Statistics, 8. Birkhäuser Boston, Inc., Boston, MA, 1985
- [6] C. Cuny, J. Dedecker and C. Jan, *Limit theorems for the left random walk on  $GL_d(\mathbb{R})$* , arXiv:1603.02217. Accepted for publication in Ann. Inst. Henri Poincaré Probab. Stat.
- [7] J. Dedecker and X. Fan, *Deviation inequalities for separately Lipschitz functionals of iterated random functions*, Stochastic. Process. Appl. 125 (2015) 3101-3132.
- [8] J. Dedecker and F. Merlevède, *Convergence rates in the law of large numbers for Banach-valued dependent random variables*, Teor. Veroyatn. Primen. 52 (2007) 562-587.
- [9] J. Dedecker, F. Merlevède, M. Peligrad and S. Utev, *Moderate deviations for stationary sequences of bounded random variables*, Ann. Inst. Henri Poincaré Probab. Stat. 45 (2009) 453-476.

- [10] A. Dembo and O. Zeitouni. *Large Deviations Techniques and Applications*, 2nd edition. Springer New York, 1998.
- [11] H. Djellout, *Moderate deviations for martingale differences and applications to  $\phi$ -mixing sequences*, Stoch. Stoch. Rep. 73 (2002) 37-63.
- [12] X. Fan, *Deviation inequalities for martingales with applications to linear regressions and weak invariance principles*, arXiv:1504.03667
- [13] M. I. Gordin, *The central limit theorem for stationary processes*, (Russian) Dokl. Akad. Nauk SSSR 188 1969 739-741.
- [14] M. I. Gordin and B. A. Lifšic, *Central limit theorem for stationary Markov processes*, (Russian) Dokl. Akad. Nauk SSSR 239 (1978), no. 4, 766-767.
- [15] S. Gouëzel and I. Melbourne, *Moment bounds and concentration inequalities for slowly mixing dynamical systems*, Electron. J. Probab. 19 (2014) 30 pp.
- [16] E. Haeusler, *An exact rate of convergence in the functional central limit theorem for special martingale difference arrays*, Z. Wahrsch. Verw. Gebiete 65 (1984) 523-534.
- [17] S. Hao and Q. Liu *Convergence rates in the law of large numbers for arrays of martingale differences*, J. Math. Anal. Appl. 417 (2014) 733-773.
- [18] C. Jan, *Vitesse de convergence dans le TCL pour des processus associés à des systèmes dynamiques ou des produits de matrices aléatoires*, Thèse de l'université de Rennes 1 (2001), thesis number 01REN10073.
- [19] E. Le Page, *Théorèmes limites pour les produits de matrices aléatoires*, Probability measures on groups (Oberwolfach, 1981), pp. 258-303, Lecture Notes in Math., 928, Springer, Berlin-New York, 1982.
- [20] Q. Liu and F. Watbled, *Exponential inequalities for martingales and asymptotic properties of the free energy of directed polymers in a random environment*, Stochastic. Process. Appl. 119 (2009) 60-90.
- [21] F. Merlevède and M. Peligrad, *Functional moderate deviations for triangular arrays and applications*, ALEA Lat. Am. J. Probab. Math. Stat. 5 (2009), 320.
- [22] A. Puhalskii, *Large deviations of semimartingales via convergence of the predictable characteristics*, Stoch. Stoch. Rep. 49 (1994), 27-85.
- [23] J. Worms, *Moderate deviations of some dependent variables. I. Martingales.*, Math. Methods Statist. 10 (2001) 38-72.
- [24] W. B. Wu and Z. Zhao, *Moderate deviations for stationary processes*, Statist. Sinica 18 (2008) 769-782.